Genet Resour Crop Evol
DOI 10.1007/s10722-017-0491-7

RESEARCH ARTICLE

Who is sowing our seeds? A systematic review of the use of plant genetic resources in research

L. R. Davies · C. J. Allender

Received: 27 July 2016 / Accepted: 16 January 2017

Abstract Collections of plant genetic resources managed by genebanks function to conserve the range of genetic diversity present in crop genepools. They can facilitate access to valuable allelic variation for both plant breeders and researchers who are able to request germplasm for use in crop improvement and both basic and applied scientific research. The direct impact of genebank collections is often unclear as downstream uses of germplasm samples may not be reported back to the genebank of origin. This study aims to systematically review scientific use of germplasm using the UK Vegetable Genebank (UKVGB) as a model. Between the years of 1980–2016, a total of 271 publications were identified as using UKVGB material. The frequency of publications and the international nature of use increased significantly over the time period studied. Accessions directly sourced from the UKVGB made up the majority of material used by researchers, but material from research-derived resources such as differential sets and core collections or diversity sets have also been used. Resistance to pests and diseases and genetic diversity were the main topics of study although germplasm was used to address a wide range of other research questions. Genebanks such as UKVGB provide an essential resource of allelic diversity in crop genepools which supports a diverse range of research projects. The utilisation of these plant genetic resources has increased over time, contributing to a substantial number of publications. Developments in sequencing technologies have no doubt played a part as larger numbers of accessions can be utilized in a single experiment, but the increase also no doubt reflects a greater interest in the use of allelic diversity to overcome challenges in crop improvement and research.

Keywords Genebank · Plant genetic resources · Research · Utilisation

Electronic supplementary material The online version of this article (doi:10.1007/s10722-017-0491-7) contains supplementary material, which is available to authorized users.

Present Address:
L. R. Davies
ADAS UK Ltd, Battlegate Road, Boxworth, Cambridge CB23 4NN, UK

L. R. Davies · C. J. Allender (✉)
Warwick Crop Centre, School of Life Sciences, University of Warwick, Wellesbourne Campus, Warwick CV35 9EF, UK
e-mail: charlotte.allender@warwick.ac.uk

Introduction

Crop domestication and breeding have resulted in varieties which display impressive adaptation to the agricultural and food systems of today. However, modern elite varieties in many crops represent only a small part of the total crop genepool when crop wild relatives and landraces are considered. This means that current breeding programmes may be lacking the allelic diversity necessary to develop the improved varieties of the future. It is therefore important to ensure the genetic diversity within the entirety of crop genepools is both conserved and available for use, by conserving the numerous varieties and landraces of each crop, as well as crop wild relatives (Jarvis et al. 2015). One effective conservation method is through ex situ storage in genebanks. Worldwide there are more than 1750 genebanks containing over 7.4 million accessions. Of these, 1240 genebanks contain crop germplasm totalling 4.6 million accessions (FAO 2010). Through conservation and subsequent distribution to plant breeders, researchers and growers, these crop genetic resources can be used to find new useful traits, such as novel sources of resistance to pests and diseases and increased tolerance to abiotic stresses. Such traits can be identified either through screening collections of both crop varieties and wild relatives (Taylor et al. 2002) or through integrating collections into differential series to produce lines that can be used in gene-for-gene models to search for specific resistance traits (Vicente et al. 2001). Crop genetic resources can also be used as tools for basic evolutionary and biological research, as well as being utilized in specific methods, such as linkage and association mapping.

The success of genebanks in collecting and conserving large-scale collections of germplasm is to be applauded, but the scale of success often means that in order to facilitate genomic and phenotyping studies, it is necessary for financial and practical reasons to develop subsets of the many thousands of accessions in crop genebank collections. These were originally defined as "core collections", designed to offer a representative sample of crop genepool variation through the selection of a limited number of accessions (Frankel et al. 1984). The core collection concept has been refined in recent years and more recent germplasm panels comprise seed collected from individual plants from each accession, reducing the potential genotypic heterogeneity of the original accession. Further refinements in highly heterozygous species can be made through the process of selfing over several generations, or the production of doubled haploid (DH) homozygous lines. Such panels of material have advantages over standard genebank accessions, in that they permit sophisticated and replicated trait screening without the confounding effects of genotypic variation among individuals. Homozygous lines are also technically easier to work with using Next Generation (NGS) sequencing and newer techniques. Core collections and derived panels of fixed diverse germplasm both offer an opportunity for researchers to pool data from different assays on the same germplasm rather than individual experiments using entirely different sets of accessions from genebank collections.

In 1980, concern over the loss of genetic diversity in vegetable crops and the potential impact this would have on global nutrition and health led to the creation of the UK Vegetable Genebank (UKVGB), currently managed by the University of Warwick. The UKVGB currently conserves approximately 14,000 accessions of crops such as alliums, brassicas, carrot, and lettuce, as well as smaller collections of minor vegetables and salad crops. Research resources such as the European Clubroot Diversity set (ECD) and the Brassica S-allele collection, are also hosted at the UKVGB. From the 1990s onwards, material from the UKVGB has been incorporated into several germplasm panels (see Table 1), and these have developed from core collections of genebank accessions to panels of homozygous lines for several different crops.

The overall aim of the UKVGB is to conserve genetic diversity within vegetable crops and their wild relatives, ensuring that the collections are available to researchers both now and in the future. The UKVGB, like other genebanks, aims to facilitate the use of germplasm by researchers, breeders and growers. In order to assist potential users, it is often helpful to know how seed requested from the collections has been used, and to be able to direct users to published datasets. In general there is a lack of information regarding the use of crop genetic resources and related outcomes (FAO 2010), with only a few published reviews, such as those by Dudnik et al. (2001), Rubenstein et al. (2006) and Dulloo et al. (2013), assessing the impact of crop genetic resources in a research context. Furthermore, two of these reviews (Dudnik et al. 2001; Dulloo et al. 2013) focused on publications from four peer-reviewed journals rather than conducting a comprehensive assessment, and therefore omitted many other publications and outputs using genebank

Table 1 Details of specific research-derived resources which include material originating from the UK Vegetable Genebank

Taxon	Resource name	Description	References
Brassica spp.	European clubroot differential series (ECD)	Differential series used to identify races of clubroot	Buczacki et al. (1975)
Brassica spp.	*Brassica* S-allele collection	*Brassica* lines with characterised S-allele haplotypes	Ockendon (2000)
Brassica oleracea L.	*Brassica oleracea* core collection (BOCC)	*B. oleracea* core collection assembled under project CT920463	Leckie et al. (1996)[a]
Brassica spp.	RESGEN *Brassica* core collections	*B. oleracea*, *B. napus* and *B. rapa* panels assembled under GENRES project CT99 109	Project respository[b]
Brassica oleracea L.	*Brassica oleracea* diversity fixed foundation set	Comprises homozygous DH or selfed lines	Walley et al. (2012)
Brassica napus L.	*Brassica napus* diversity fixed foundation set	Comprises homozygous DH or selfed lines	Walley et al. (2012)
Brassica spp.	Brassica C-genome diversity fixed foundation set	Comprises DH lines derived from cross between wild species and a rapid cycling line	Walley et al. (2012)
Brassica napus L.	ASSYST		Bus et al. (2011)
Daucus carota L.	Carrot diversity set (CDS)	Set of UKVGB accessions and other research lines developed by VeGIN[c] project	Project website[d]
Allium cepa L.	Onion diversity set	Set of UKVGB accessions developed by VeGIN project	Project website[d]
Lactuca sativa L.	Lettuce diversity set	Set of genebank accessions including UKVGB material developed by VeGIN project	Project website[d]

[a] Project repository available at http://ecpgr.cgn.wur.nl/Brasedb/brasair.htm

[b] Project repository available at http://ecpgr.cgn.wur.nl/Brasedb/brasresgen.htm

[c] Vegetable Genetic Improvement Network

[d] Project website at vegin.warwick.ac.uk

material. An overview of the activities of the genebank at IPK Gatersleben, Germany was provided by Hammer et al. (1994) in a description of the achievements of the genebank over a fifty year period. Understanding the impact of genebank collections is a prerequisite for making the case for continuing operations and funding. The aim of this study is to systematically investigate the research outputs associated with a single collection of plant genetic resources (the UK Vegetable Genebank) through a search for relevant published scientific literature. This will provide an in-depth analysis of the nature of research questions addressed, the geographic location of authors and users and the types of material most frequently utilised. The collation of all identifiable publications relating to the UKVGB collection will also aid future users through the clear identification of pre-exiting datasets related to the collections, assisting in the selection of appropriate material and reducing the need to duplicate work.

Materials and methods

Search method

A systematic survey of scientific literature was carried out during November 2015. The search was complicated by the fact that since its creation, the UKVGB has undergone a number of name changes. Originally part of the National Vegetable Research Station (NVRS), in 1990 a merger between NVRS and a number of other UK horticulture research institutions led to the creation of Horticulture Research International (HRI). Subsequently, in 2004 HRI, Wellesbourne, was incorporated into the University of Warwick creating Warwick HRI. At this time the genebank became known as Warwick GRU. In 2012 the School of Life Sciences was created and the genebank is now known as the UKVGB. Therefore, due to these name changes it was necessary to search for references to all previous names of the UKVGB and acronyms of those names.

Searches of databases and websites were made in November 2015, with further enquiries being made up until January 2016. Online sources were searched from 1980 to the present, covering the period that the UKVGB has been in operation. Details of the search engines and search terms used can be found in ESM 1 (Appendix 1). As the nature of the search was for specific reference of the UKVGB (including previous names) in the body of the text of publications, only search engines that searched the entire article in addition to title and abstract were used.

Papers provided by colleagues known to have used UKVGB material were also included. The historical records available at the UKVGB were also searched for publications. References in the selected publications were searched for any other publications that may have used UKVGB material. Where publications produced new experimental lines and resources using UKVGB material, publications citing these were also searched and included if they used these lines or collections derived from UKVGB material.

Inclusion criteria

During the initial stages of the search, papers were either excluded as not being relevant based on the title or abstract or, if potentially relevant, the text was searched for mention of NVRS, HRI, HRIGRU, Warwick GRU, UKVGB, or Warwick Crop Centre (see ESM 1 for explanation of acronyms). Papers were then excluded as having no mention of these terms, or if they did not describe the use of plant germplasm. After the conclusion of the search, a list of 379 publications which potentially used material from UKVGB was assembled. These publications were then sorted into two categories: those that had clearly used germplasm from UKVGB and those where usage was unclear. Details of this search can be found in ESM 1. The Materials and methods, results, acknowledgements, and where possible supplementary data sections of these publications were thoroughly searched for mention of material from UKVGB, or previous names. Publications that did not use UKVGB material, but had used other unrelated germplasm resources were then excluded.

After this search, publications in which use of UKVGB material was still uncertain due to the use of non-matching accession identifiers or a lack of information on seed provenance, were further investigated either by examining the accessions used in the publications and comparing these accessions to the UKVGB database records, or further examining the source of the material. Papers which used research or breeding lines that were sourced from NVRS/HRI, but did not originate from UKVGB were excluded, as well as papers where accession names and numbers did not match passport data in the UKVGB database. For studies published after 2007 where it was still unclear if UKVGB material was used, the corresponding author was contacted to request further information. The search process is represented in the flowchart in ESM 1 (Appendix 2).

Research-derived resources

Seed from the UKVGB has been incorporated into a number of core collections and diversity sets (Table 1), such as the *Brassica oleracea* Diversity Fixed Foundation Set (BolDFFS—see Walley et al. 2012) where fixed lines were produced, mainly from a *B. oleracea* L. core collection created by Leckie et al. (1996). UKVGB *Brassica* accessions have also incorporated into core collections generated in the RESGEN project, where collections of *B. oleracea*, *B. napus* L., *B. rapa* L. and *B. carinata* A. Braun were created to increase knowledge about the genetic resources available in these species and enable these resources to be used by growers and breeders (Lühs et al. 2003). Subsequently, UKVGB accessions were used to generate inbred lines included in the ASSYST *B. napus* diversity set (pers. comms. R.J. Snowdon, Bus et al. 2011). UKVGB accessions have also been incorporated into a lettuce diversity set (Burns et al. 2011), carrot diversity set (see Table 1), onion diversity set (see Table 1), and *B. napus* diversity set (Taylor et al. 2015). Specific searches for publications using these research-derived resources were undertaken, and added to the group of publications identified previously.

Data extraction

The following data were extracted from all publications using UKVGB material and recorded: publication type, year of publication, institution and country of all authors, taxonomic identity of germplasm used, accession identifiers and research topic. For multi-author, multi-institution publications, the publication was counted once for each institution. Likewise, where authors from more than one country contributed to a publication, the publication counted once for each country. The same approach was taken with regard to the taxonomic identity of the germplasm used and the topic of study, where for example more than one pest or pathogen was the subject of the study. R (version 3.2.2) was used to conduct a linear regression to assess the change in number of publications over time and the change in the number of countries with contributing authors over time.

Results

Publications identified

A total of 271 publications were found to have used material from the UKVGB during the period 1980-January 2016. Of the 271 publications that used UKVGB material 218 (80.4%) were published in peer reviewed journals, 16 (5.9%) were theses (both MSc (2) and PhD (14)), 14 (5.2%) were published in non-peer reviewed journals, 8 (3%) were conference papers, 6 (2.2%) were final reports for funding bodies, 3 (1.1%) were published in books, 2 (0.74%) were working group reports, 2 (0.74%) were patents, 1 (0.37%) was of unknown publication status, and 1 (0.37%) was a newsletter. A total of 88 peer-reviewed journals published papers using UKVGB materials. Of these, 5 journals published 67 of the articles, with Theoretical Applied Genetics publishing 20 papers, Euphytica publishing 16 papers, Genetic Resources and Crop Evolution publishing 13 papers, Plant pathology publishing 10 papers, and Annals of Applied Biology publishing 8 papers. The remaining 151 papers were published in 83 other diverse journals.

Annual frequency of publications

In the years between 1980 (when UKVGB was opened) and 2015 the number of publications increased, from 0 for the period 1980–1985 to 73 for the period 2011–2015, with a mean number of publications of 14.6 ($\pm$1.03) per year. Linear regression showed there was a significant ($p < 0.001$, adj R square $= 0.819$) increase in the number of publications published per year between the years of 1980–2015 (Fig. 1).

Where is the material being used?

Authors of the identified publications were based in 189 organisations located in 36 countries. Of these institutions, a total of 120 are located in Europe, 28 in Asia, 24 in North America, 12 in Oceania, 3 in South America, and 2 in Africa. The institution with the most publications was HRI with 44 publications, followed by the University of Warwick with 35 publications (although it should be noted that HRI became part of the University of Warwick in 2004). Of the 11 institutions that published the most papers 3 are based in the UK, 2 in the USA, 2 in Canada, 1 in Germany, 1 in the Netherlands, 1 in Portugal and 1 in France.

Linear regression showed that the international diversity of author institutional affiliation significantly increased over the period studied ($p < 0.001$, adj R square $= 0.609$, data not shown) with 1 country in 1986 (the year of the first publication using UKVGB material) to 9 in 2015, with a peak of 18 countries in 2013. A total of 17 countries are in Europe, 10 in Asia, 3 in South America, 2 in North America, 2 in Oceania, and 2 in Africa. The UK was the country with the most publications with 107 (29.8%), followed by the USA with 44 (12.3%). Since 2000, there has been a large increase of the number of publications produced by authors based in institutions located in Asia.

What material is being used?

Across the 271 publications, accession identifiers were provided for 1282 individual accessions, including those used in diversity sets. A further 15 lines from the ECD Series, and 47 Brassica S-allele lines, were also used. UKVGB accessions not incorporated in diversity sets were used in the majority of publications, followed by the S-allele collection, and the ECD Series (Table 2). Accessions from 7 diversity sets were used in a total of 47 publications, with the number of publications using diversity sets increasing from 1 in 1997 to 6 in 2015.

A total of 172 taxa from five plant families were accessed from the UKVGB collections [Apiaceae (63), Amaryllidaceae (21), Asteracaea (10), Brassicaceae (77), and Fabiacaea (1)]. The UKVGB routinely classifies accessions to species or subspecies level. The most frequently used species overall was *B. oleracea* with 437 occurrences of all *B. oleracea*

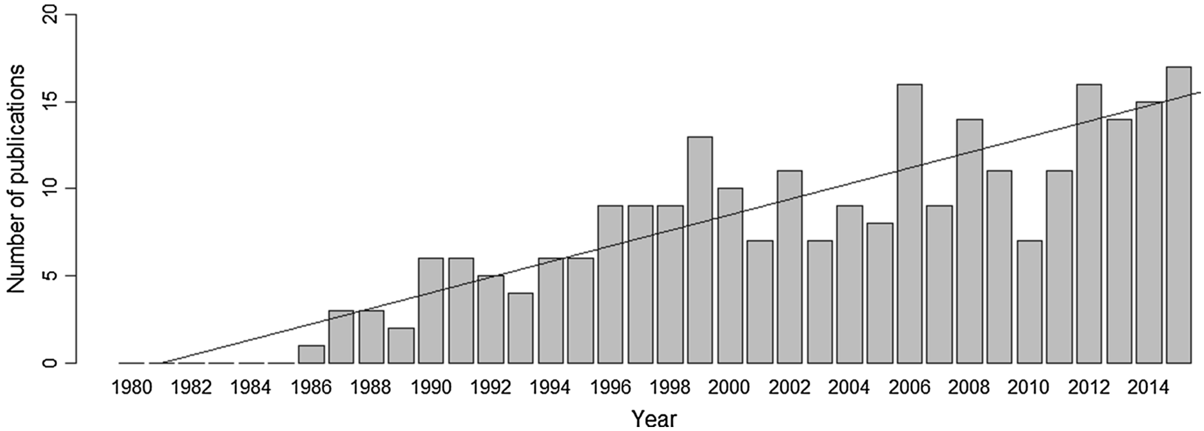

Fig. 1 Number of publications published per year between the years of 1980–2015 (excluding 2016) using UKVGB material. Linear regression ($y = 0.446x - 883.7$) shows a significant increase ($p < 0.001$, adj R square -0.819) in the number of publications over time

Table 2 Type of material from UKVGB used in publications, including genetic diversity sets

UK Vegetable Genebank (UKVGB) material	Number of publications
UKVGB material (HRIGRU) not accessed via diversity sets	178
S-allele collection	22
ECD series	17
ASSYST diversity set	12
Brassica oleracea Core Collection	12
Brassica oleracea Diversity Foundation Set	9
UKVGB material incorporated in other genebanks	5
Carrot Diversity Set	5
RESGEN	6
Lettuce Diversity Set	4
Brassica napus Diversity Foundation Set	1

subtaxa in the publications identified during this review.

Topics researched using plant genetic resources from the UKVGB

A total of 21 main subject areas were identified for the 271 publications (Table 3). Resistance to pests and disease was the most common subject area, followed by broad investigations of genetic diversity, and pest biology studies, where UKVGB accessions were used as host plants.

Of the publications that investigated resistance to pests and diseases, oomycetes and insect pests were the pest groups most investigated. UKVGB accessions were used to investigate resistance and susceptibility to 24 different species. *Brevicoryne brassicae* (cabbage aphid) was the species most investigated in terms of research into host plant resistance (8 papers), followed by *Xanthamonus campestris* pv. *campestris* (Pammel) Dowson (blackrot-6 papers), *Albugo candida* (Pers.) Kuntze (white rust-5 papers), *Hyaloperonospora* Gaum (downy mildew-5 papers), and *Plasmodiophora brassicae* Woronin (clubroot-5 papers). The remaining 19 species were studied in a total of 34 publications.

Where UKVGB accessions were used as host plants, *P. brassicae* (a Phytomyxea and the causal agent of clubroot) was the pathogen most studied, being reported in 16 publications and reflecting the uptake and usage of the ECD Series. The biology of a further 16 plant pests and pathogens was also investigated, including 4 studies on *Psila rosae* F. (carrot fly) and 2 on *A. candida*. For the remaining 14 species

Table 3 Subject area of the main objective of the 271 publications using UKVGB material

Subject	Number of papers
Resistance to pests and disease	55
Genetic diversity/variation	43
Pest biology (microbial and animal)	35
Taxonomy and phylogeny	24
Self-incompatibility	22
Secondary metabolites	14
Genetic mapping	13
Flowering	13
Breeding/hybridization	9
Nutrient uptake	7
Genetic resources unit/core collection	7
Seed/seedling/plant development	6
Fungal interactions	4
Bulb biology	3
Crop wild relatives	3
Cultivation	3
Genetic biofortification	3
Climate change	2
Food safety	2
Gene flow	2
Transgenes	1

only one publication studied each of the species. *A. candida*, *Bremia lactucae* Regel, *B. brassicae*, *P. brassicae*, *Sclerotinia sclerotiorum* (Lib.) de Barry, and *X. campestris* pv. *campestris* were species where UKVGB material was used both as host plant material to study the diseases and as a source of material to investigate resistance to the diseases.

Discussion

Underlying importance of plant genetic resources in a research context

The value of collections of plant genetic resources to scientific research is clear from the 271 publications identified through the systematic review; this reflects only the usage of the UKVGB collections, which is limited by its remit to vegetable crops and does not cover the more intensively researched cereal crops. The majority of scientific outputs have been in peer-reviewed journals, with 5 journals publishing a total of 63 of the articles. Two of these journals, Euphytica and Theoretical and Applied Genetics, were journals that Dudnik et al. (2001) and Dulloo et al. (2013) searched to evaluate the patterns of use of plant genetic resources, however, the total number of peer-reviewed journals that published papers using UKVGB material highlights the need for broad searches of the literature, particularly when assessing the usage of germplasm from a single genebank.

The use of UKVGB material has increased significantly over time, with 2015 producing the most publications to date. Some of this increase can be attributed to more stringent journal standards over the time period studied, leading to better reporting of the provenance of plant germplasm in more recent studies. It is likely that publications were missed, particularly toward the start of the period in question simply because authors did not clearly detail the source of their experimental material. In addition, some of the papers identified in this study indicated UKVGB as a germplasm source but did not specify the individual accessions used. However, most of the increase in usage is likely to be the result of increased access to UKVGB resources by researchers. This is a trend reported across other genebanks (Dulloo et al. 2013). The differences in number of publications year to year, with an overall increasing trend, follows the sawtooth pattern of genebank accession requests reported by Widrlechner and Burke (2003), where periods of high use precede and follow periods of low use, but with an overall increasing or decreasing trend over a 5-year period.

Global impact of the UKVGB collections

Not only is the use of material from the UKVGB increasing, the collections are of increasing global significance, with the number of countries where authors were based increasing significantly over time. Institutions in Europe were the major users of UKVGB material, but institutions from all continents used UKVGB material, with an increasing use by authors based in institutes in Asia over the past 15 years. HRI Wellesbourne and the University of Warwick were the two institutions that produced the most publications using UKVGB material. Historically the major focus of vegetable research in the UK was concentrated at the research institute which began as NVRS and became HRI and latterly merged with the University of Warwick. This was a major factor in the location of the genebank, enabling it to be better integrated with the UK vegetable research community. Therefore, considering the focus of the research institute based at Wellesbourne and the access its researchers have to the UKVGB, it is no surprise that HRI Wellesbourne and the University of Warwick are the main users of the UKVGB. Furthermore, the other main institutes using UKVGB material also focus on crop protection, for example, the University of Wisconsin-Madison has a vegetable research station, as well as, a plant breeding and plant genetics programme, and one of the main focuses of the Instituto Superior de Agronomia is plant protection, again making it unsurprising that these institutes are major historical users of UKVGB material.

What is UKVGB material being used for?

Year on year average crop yield losses due to animal pests, viruses, and pathogens are 10, 4, and 10%, respectively (Oerke 2006), although these losses can vary massively with pest and disease outbreaks (Strange and Scott 2005). Current control methods have little effect on reducing yield losses due to these pests and therefore understanding pest and disease biology and the identification of new sources of resistance are significant research areas in crop protection (Strange and Scott 2005), and it is no surprise that these were two of the most widely investigated subject areas in publications using UKVGB material. This use of material for pest and disease research is also reflected in the use of other genebank collections, such as those stored by the Genetic Resources Information Network of the United States Department of Agriculture (as noted by Rubenstein et al. 2006).

Investigating crop genepool diversity was another major topic of the publications identified. The crop genetic resources stored at UKVGB have enabled researchers to further study genetic diversity and integrate genetic resources from crop wild relatives and multiple varieties of crops into their studies. As genetic diversity is relatively low among crop species (Warchefsky et al. 2014) it is not only essential to maintain this diversity, it is important to quantify the diversity is present. This information can then be used in new breeding technologies, such as next generation sequencing (Walley et al. 2012), and to identify new resources for improving varieties (Rubenstein et al. 2006), helping to improve food security and nutrition. Information on crop genetic diversity can also be used by plant breeders and genebank curators when propagating or maintaining lines, as it can provide information on the number of individuals needed to maintain genetic diversity within a line and prevent inbreeding depression (Baldwin et al. 2012).

The impact of refined and research-derived subsets of germplasm

Characterised crop genetic diversity has been refined and is available to users in a range of different forms, including differential series, set of lines with defined haplotypes (such as the *Brassica* S-allele collection) and core collections and diversity sets. The latter two offer users access to a good representation of crop genepool diversity in a manageable number of accessions or lines. One may expect that the use of diversity sets/core collections would increase over time. A total of 87 publications cited the use of research-derived resources (Table 2). The well-established resources such as the ECD Series and the *Brassica* S-allele collection were understandably cited the most frequently, although the newly developed diversity sets for *B. oleracea* and the *B. napus* ASSYST set were cited in 12 publications. Core collections such as those created as part of specific projects (EU AIR and RESGEN) seem to reach a peak of citations shortly after the conclusion of the project (presumably most of these report directly on the project's activities). The impact diversity sets had on the number of publications using UKVGB material is lower than may be expected considering the number and prominence of these sets (Walley et al. 2012). However, this study covers a period of 35 years and diversity sets are a more recent development, with the *B. oleracea* core collection created in 1996 (Leckie et al. 1996), being replaced with the *B. oleracea* diversity foundation set in 2008 (Pink et al. 2008; Walley et al. 2012), and the ASSYST diversity set created as recently as 2011 (Bus et al. 2011). This is in contrast to the ECD and S-allele sets, which were used in more publications, but were created before the UKVGB opened and were subsequently integrated into the UKVGB collection (Dixon 2009; Ockendon 2000). The use of diversity sets is gradually increasing, from 1 publication in 1997 to 6 in 2015, suggesting that diversity sets are having an increasing importance in research. Nevertheless, there is a major underlying use of material directly sourced from the UKVGB (and not via the maintainers of diversity sets) with the majority of publications using this material. This indicates the importance of the conservation of entire collections is still very much required, and that diversity sets and similar refined resources complement genebank collections rather than replacing them.

The importance of specific research resources is highlighted by the fact that the ECD and S-allele collections stored at UKVGB have been cited in many publications. Recently, the ECD collection has been utilised by researchers in Canada where clubroot (*P. brassicae*) is becoming a major problem in brassica crops and spreading quickly across the country (Cao et al. 2009). In infected fields, clubroot causes an average yield loss of 11% (Dixon 2009), with some strains infecting 74–100% of a crop (Cao et al. 2009). The ECD collection is enabling researchers in Canada and elsewhere to characterise the spread of the disease and search for new sources of resistance (Strelkov et al. 2006), showing the importance of maintaining research-derived resources like the ECD Series and ensuring their future availability for research and breeding.

The impacts of model species and developments in sequencing technologies

Accessions of Brassicaceae species were used in the majority of publications, with this use reflecting the accessions conserved at the UKVGB, where >54% of the species concerned are Brassicaceae. The use of Brassicaceae species in publications will also be driven by the fact that brassica crops are one of the most widely grown vegetable crops in addition to the close taxonomic relationship between Brassica species and the model plant *Arabidopsis thaliana* L. The latter offers researchers the opportunity to translate genetic and genomic studies from model species to crops (Parkin 2011). In contrast, economically significant crops such as alliums may be less frequently targeted by genetic/genomic studies due to high levels of duplication and low gene densities (King et al. 1998; Jakše et al. 2008). The significant decrease in costs associated with whole genome sequencing and high throughput genotyping mean that it is now feasible for researchers to study larger numbers of accessions.

Conclusion

This report has highlighted the importance of the role played by collections of plant genetic resources in supporting a broad range of crop and plant science. Considering UKVGB is a relatively small genebank with a specific remit for vegetable crops, it is remarkable that 271 publications indicated the use of germplasm from UKVGB. This number is very likely to be an under-estimate due to under citation of the source of germplasm, especially during the early period of operation of the genebank. The development of refined panels of germplasm in the form of diversity sets, based on core collections is relatively new; it is difficult at the present time to determine their impact on the scientific literature compared to standard genebank accessions. The continued use of both genebank accessions and diversity sets indicates the importance retaining secure access to both in the future.

Acknowledgements The UK Vegetable Genebank is funded by the Department for the Environment, Food and Rural Affairs under project GC0148. The systematic review was supported by the University of Warwick.

Compliance with ethical standards

Conflicts of interest C. Allender has led the UK Vegetable Genebank project since 2014.

References

Baldwin S, Pither-Joyce M, Wright K, Chen L, McCallum J (2012) Development of robust genomic simple sequence repeat markers for estimation of genetic diversity within and among bulb onion (*Allium cepa* L.) populations. Mol Breed 30:1401–1411

Buczacki ST, Toxopeus H, Mattusch P, Johnston TD, Dixon GR, Hobolth LA (1975) Study of physiologic specialization in *Plasmodiophora brassicae*: proposals for attempted rationalization through an international approach. Trans Br Mycol Soc 65(2):295–303

Burns IG, Zhang K, Turner MK, Lynn J, McClement S, Hand P, Pink D (2011) Genotype and environment effects on nitrate accumulation in a diversity set of lettuce accessions at commercial maturity: the influence of nitrate uptake and assimilation, osmotic interactions and shoot weight and development. J Sci Food Agric 91:2217–2233

Bus A, Körber N, Snowdon RJ, Stich B (2011) Patterns of molecular variation in a species-wide germplasm set of *Brassica napus*. Theor Appl Genet 123:1413–1423

Cao T, Manolii VP, Hwang S-F, Howard RJ, Strelkov E (2009) Virulence and spread of *Plasmodiophora brassicae* (clubroot) in Alberta, Canada. Can J Plant Pathol 31:321–329

Dixon GR (2009) The occurrence and economic impact of *Plasmodiophora brassicae* and clubroot disease. J Plant Growth Regul 28:194–202

Dudnik NS, Thormann I, Hodgkin T (2001) The extent of use of plant genetic resources in research—a literature survey. Crop Sci 41:6–10

Dulloo ME, Thormann I, Fiorino E, De Felice S, Rao VR, Snook L (2013) Trends in research using plant genetic resources from germplasm collections: from 1996 to 2006. Crop Sci 53:1217–1227

FAO (2010) The second report on the state of the world's plant genetic resources for food and agriculture. FAO, Rome

Frankel OH et al (1984) Genetic perspectives of germplasm conservation. In: Arber WK et al (eds) Genetic manipulation: impact on man and society. Cambridge University Press, Cambridge, pp 161–170

Hammer K, Gäde H, Knüpffer H (1994) 50 Jahre Genbank Gatersleben–eine Übersicht. Vortr. Pflanzenzüchtung 27:333–383

Jakše J, Meyer JDF, Suzuki G, McCallum J, Cheung F, Town CD, Havey MJ (2008) Pilot sequencing of onion genomic DNA reveals fragments of transposable elements, low genetic densities, and significant gene enrichment after methyl filtration. Mol Genet Genomics 280:287–292

Jarvis S, Fielder H, Hopkins J, Maxred N, Smart S (2015) Distribution of crop wild relatives of conservation priority in the UK landscape. Biol Conserv 191:444–451

King JJ, Bradeen JM, Bark O, McCallum JA, Havey MJ (1998) A low-density genetic map of onion reveals a role for tandem duplication in the evolution of an extremely large diploid genome. Theor Appl Genet 96:52–62

Leckie D, Astley D, Crute IR, Ellis PR, Pink DAC, Boukema I, Monterio AA, Dias S (1996) The location and exploration of genes for pest and disease resistance in European gene bank collections of horticultural Brassicas. Acta Hort 407:95–101

Lühs W, Seyis F, Frauen M, Busch H, Frese L, Willner E, Friedt W, Gustafsson M, Poulsen G (2003) Development and evaluation of a *Brassica napus* core collection. In: Knüpffer H, Ochsmann J (eds) Rudolf Mansfeld and plant genetic resources. Proceedings of a symposium dedicated to the 100th birthday of Rudolf Mansfeld, Gatersleben, Germany, 8–9 October 2001. Schriften zu Genetischen Ressourcen 19, pp. 287–292. ZADI/IBV, Bonn

Ockendon DJ (2000) The S-allele collection of *Brassica oleracea*. Acta Hort 539:25–30

Oerke E-C (2006) Crop losses to pests. J Agr Sci 144:31–43

Parkin I (2011) Chapter 5: Chasing ghosts: comparative mapping in Brassicaceae. In: Schmidt R, Bancroft I (eds) Genetics and genomics of the Brassicaceae, plant 153 genetics and genomics: crops and models 9. Springer, Berlin

Pink DAC, Bailey L, McClement S, Hand P, Mathas E, Buchanan-Wollaston V, Astley D, King GJ, Teakle G (2008) Double haploids, markers and QTL analysis in vegetable brassicas. Euphytica 164:509–514

Rubenstein KD, Smale M, Widrlechner MP (2006) Demand for genetic resources and the U.S. national plant germplasm system. Crop Sci 46:1021–1031

Strange RN, Scott PR (2005) Plant disease: a threat to global food security. Annu Rev Phytopathol 43:83–116

Strelkov SE, Tewari JP, Smith-Degenhardt E (2006) Characterization of *Plasmodiophora brassicae* populations from Alberta, Canada. Can J Plant Pathol 28:467–474

Taylor JD, Conway J, Roberts SJ, Astley D, Vicente JG (2002) Sources and origin of resistance to *Xanthomonas campestris* pv. *campestris* in *Brassica* genomes. Am Phytopathol Soc 92:105–111

Taylor A, Coventry E, Jones JE, Clarkson JP (2015) Resistance to a highly aggressive isolate of *Sclerotinia sclerotiorum* in a *Brassica napus* diversity set. Plant Pathol 64:932–940

Vicente JG, Conway J, Roberts SJ, Taylor JD (2001) Identification and origin of *Xanthomonas campestris* pv. *campestris* races and related pathovars. Am Phytopathol Soc 91:492–499

Walley PG, Teakle GR, Moore JD, Allender CJ, Pink DAC, Buchanan-Wollaston V, Barker GC (2012) Developing genetic resources for pre-breeding in *Brassica oleracea* L.: an overview of the UK perspective. J Plant Biotechnol. 39:62–68

Warchefsky E, Penmetsa RV, Cook DR, von Wettberg EJB (2014) Back to the wilds: tapping evolutionary adaptations for resilient crops through systematic hybrisization with crop wild relatives. Am J Bot 101(10):1791–1800

Widrlechner MP, Burke LA (2003) Analysis of germplasm distribution patterns for collections held at the North Central Regional Plant Introduction Station, Ames, Iowa, USA. Genet Resour Crop Evol 50:329–337